

## CLAIMS

What is claimed is:

- 1 1. A method for recognizing speech, the method comprising:  
2 recognizing a sequence of words;  
3 processing the sequence of words using word agglomeration; and  
4 classifying the processed sequence of words as a predetermined command.
- 1 2. The method of claim 1, further comprising performing an action corresponding to  
2 the predetermined command.
- 1 3. The method of claim 1, wherein processing comprises replacing the sequence of  
2 words with an associated word  $n$ -tuple sequence.
- 1 4. The method of claim 1, wherein the associated word  $n$ -tuple sequence is a  
2 sequence of all strings of  $n$  consecutive words present in the sequence of words.
- 1 5. The method of claim 3, wherein classifying comprises semantically inferring the  
2 predetermined command from the associated word  $n$ -tuple sequence.
- 1 6. The method of claim 1, wherein classifying comprises semantically inferring the  
2 predetermined command from the processed sequence of words.

1 7. The method of claim 6, wherein semantically inferring comprises determining a  
2 correlation between the processed sequence of words and at least one semantic anchor.

1 8. The method of claim 7, wherein the correlation is a distance between a vector  
2 corresponding to the processed sequence of words and a vector corresponding to the at  
3 least one semantic anchor.

1 9. The method of claim 8, wherein semantically inferring further comprises selecting  
2 the predetermined command from the semantic anchor vector having the shortest  
3 distance.

1 10. The method of claim 9, wherein the semantic anchor represents a one of a  
2 plurality of predetermined commands.

1 11. The method of claim 7, wherein the at least one semantic anchor is derived from a  
2 training data.

1 12. The method of claim 6, wherein semantically inferring the predetermined  
2 command depends on the order of the words in the processed sequence of words.

1 13. The method of claim 1, wherein the classifying comprises:  
2 generating a vector representation of the processed sequence of words; and

3 comparing the vector representation to a plurality of semantic anchors, wherein  
4 each of the plurality of semantic anchors corresponds to one of a plurality of voice  
5 commands.

1 14. The method of claim 13, wherein the classifying further comprises:  
2 selecting a semantic anchor of the plurality of semantic anchors that is most  
3 similar to the vector representation; and  
4 classifying the processed sequence of words as the command that corresponds to  
5 the selected semantic anchor.

1 15. The method of claim 14, wherein the selecting comprises:  
2 for each of the plurality of semantic anchors, identifying the similarity between  
3 the vector representation and the semantic anchor by calculating the cosine of the angle  
4 between the product of the vector representation and a diagonal matrix of singular values  
5 and the product of the semantic anchor and the diagonal matrix of singular values; and  
6 selecting the semantic anchor of the plurality of semantic anchors that  
7 corresponds to the largest cosine value as the semantic anchor that is most similar to the  
8 vector representation.

1 16. The method of claim 13, wherein the vector representation is an indication of how  
2 frequently each of a plurality of word  $n$ -tuples occurs within the processed sequence of  
3 words.

1 17. The method of claim 16, wherein each of the plurality of semantic anchors is an  
2 indication of how frequently each of the plurality of word  $n$ -tuples occurs with respect to  
3 the corresponding command.

1 18. The method of claim 13, wherein each of the plurality of semantic anchors  
2 represents a plurality of different ways of speaking the corresponding command.

1 19. The method of claim 13, wherein each of the plurality of semantic anchors  
2 represents a plurality of different commands having the same words, but in a different  
3 order.

1 20. A machine-readable medium having stored thereon a plurality of instructions that,  
2 when executed by a processor, cause the processor to recognize a voice command by:  
3 recognizing a sequence of words;  
4 processing the sequence of words using word agglomeration; and  
5 classifying the processed sequence of words as a predetermined command.

1 21. The machine-readable medium of claim 20, wherein the plurality of instructions  
2 further cause the processor to perform an action corresponding to the predetermined  
3 command.

1 22. The machine-readable medium of claim 20, wherein the plurality of instructions  
2 for processing comprises:

3 replacing the sequence of words with an associated word  $n$ -tuple sequence.

1 23. The machine-readable medium of claim 22, wherein the associated word  $n$ -tuple  
2 sequence is a sequence of all strings of  $n$  consecutive words present in the sequence of  
3 words.

1 24. The machine-readable medium of claim 22, wherein the instructions for  
2 classifying comprises semantically inferring the predetermined command from the  
3 associated word  $n$ -tuple sequence.

1 25. The machine-readable medium of claim 22, wherein the instructions for  
2 semantically inferring comprises determining a correlation between a semantic  
3 representation of the associated word  $n$ -tuple sequence and at least one semantic anchor.

1 26. The machine-readable medium of claim 25, wherein the instructions for  
2 determining a correlation comprise determining a distance between a vector  
3 corresponding to the semantic representation and a vector corresponding to the at least  
4 one semantic anchor.

1 27. The machine-readable medium of claim 26, wherein the instructions for  
2 semantically inferring further comprises selecting the predetermined command from the  
3 semantic anchor vector having the shortest distance to the a vector corresponding to the  
4 semantic representation.

1 28. The machine-readable medium of claim 25, wherein the at least one semantic  
2 anchor represents a one of a plurality of predetermined commands.

1 29. The machine-readable medium of claim 25, wherein the at least one semantic  
2 anchor is derived from a training data.

1 30. The machine-readable medium of claim 25, wherein the performance of the  
2 instructions for semantically inferring the predetermined command depends on the order  
3 of the words in the processed sequence of words.

1 31. An apparatus for recognizing a voice command, the apparatus comprising:  
2 a speech recognizer to recognize a sequence of words received as the voice  
3 command;  
4 a processor to process the sequence of words using word agglomeration; and  
5 a semantic classifier, coupled to the processor, to semantically infer from a vector  
6 representation of the processed sequence of words which of a plurality of predetermined  
7 commands correlate to the voice command.

1 32. The apparatus of claim 31, further comprising:  
2 an action generator, coupled to the semantic classifier, to use the vector  
3 representation to determine an action to be performed.

1 33. The apparatus of claim 31, wherein the semantic classifier is further to compare  
2 the vector representation to a plurality of semantic anchors, wherein each of the plurality  
3 of semantic anchors corresponds to a one of the plurality of predetermined commands.

1 34. The apparatus of claim 33, wherein the semantic classifier is further to identify a  
2 semantic anchor of the plurality of semantic anchors that is most similar to the vector  
3 representation, and to classify the vector representation as the one of the plurality of  
4 predetermined commands that corresponds to the identified semantic anchor.

1 35. An apparatus for recognizing a voice command, the apparatus comprising:  
2 means for recognizing a sequence of words received as the voice command;  
3 means for processing the sequence of words using word agglomeration; and  
4 means, coupled to the means for processing, for semantically inferring from a  
5 vector representation of the processed sequence of words which of a plurality of  
6 predetermined commands correlate to the voice command.

1 36. The apparatus of claim 35, further comprising means, coupled to the means for  
2 semantically inferring, for using the vector representation to determine an action to be  
3 performed.







